



## King's Research Portal

### *Document Version*

Early version, also known as pre-print

[Link to publication record in King's Research Portal](#)

### *Citation for published version (APA):*

Brewer, B. (1995). Mental Causation: Compulsion by Reason. *Proceedings of the Aristotelian Society Supplementary Volumes*, 69, 237.

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



**Open Access document  
downloaded from King's Research Portal  
<https://kclpure.kcl.ac.uk/portal>**

**Citation to published version:**

Brewer, B. (1995). Mental Causation: Compulsion by Reason. Proceedings of the Aristotelian Society Supplementary Volumes, 69, 237.

**The published version is available at:**

DOI: N/A

**This version:** Preprint

URL identifying the publication in the King's Portal:

<https://kclpure.kcl.ac.uk/portal/en/publications/mental-causation-compulsion-by-reason%28d5d57c4d-c168-471f-9782-d719832370b4%29.html>

URL of this Open Access document in the King's Portal:

<https://kclpure.kcl.ac.uk/portal/files/5966709/mccr.pdf>

**The copyright in the published version resides with the publisher.**

**When referring to this paper, please check the page numbers in the published version and cite these.**

**General rights**

Copyright and moral rights for the publications made accessible in King's Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications in King's Research Portal that users recognise and abide by the legal requirements associated with these rights.'

- Users may download and print one copy of any publication from King's Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the King's Research Portal

**Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# *Mental Causation: Compulsion by Reason*

BILL BREWER

## 1. Introduction

The standard paradigm for mental causation is a person's acting for a reason. Something happens - she intentionally  $\phi$ 's - the occurrence of which we explain by citing a relevant belief or desire. In the present context, I simply take for granted the following two conditions on the appropriateness of this explanation. First, the agent  $\phi$ 's *because* she believes/desires what we say she does, where this is expressive of a *causal* dependence.<sup>1</sup> Second, her believing/desiring this gives her a *reason* for  $\phi$ -ing: recognizing that she has this belief/desire makes her  $\phi$ -ing intelligible as rational in the light of her other attitudes and circumstances. A further condition must be met, though, if this is to be a genuine psychological explanation, a case of her acting *for* the reason in question. Consider the following example of Davidson's (1973, p. 79). An exhausted climber is desperate to rid herself of the weight and danger of holding her partner on a rope; and her sudden realization that simply letting go would achieve this so unnerves her that her grip loosens slightly and he falls. Her releasing him causally depends upon her having this belief and desire, which provide *a* reason for doing what she does. But this is not *why* she does it: it would be at best misleading to say that she dropped him, intentionally, because she was fed up with holding his weight, or because she thought that she might otherwise fall. Her letting go does not depend upon her having these reasons in the right way. The reason-giving relation is causally irrelevant. If we are to explain a person's acting *for* a reason, then her doing

---

<sup>1</sup> For argument in favour of this assumption, see Davidson (1963) and Child (1994, ch. 3). For dissension, see Taylor (1964, ch. 1) and Stoutland (1976).

what she does must causally depend upon her having the cited reason *in virtue of* its rationalizing an action of the relevant type.<sup>2</sup>

I shall be concerned primarily in what follows with precisely this notion of causal dependence in virtue of rationalization as it figures in equally important but less widely discussed cases of mental causation. Two phenomena I shall be particularly interested in are the following.

1. The dependence of a person's correct application of a linguistic expression, in a judgement expressive of knowledge, upon his understanding of the expression in question.
2. The dependence of a person's acquisition of knowledge through deductive argument upon his grasp of the rules of inference governing the constituent logical constants.

Each of these concerns the exercise of a cognitive capacity: linguistic understanding and deductive reasoning. Acquiring such a capacity involves attaining some conception of its correct use; exercising it involves making an appropriate judgement in the light of this conception. In the linguistic case, learning what particular words or expressions mean involves attaining some appreciation of what count as their correct and incorrect future applications; using them knowledgeably in the categorization of objects in thought or communication is an exercise of this appreciation of the norms governing them. Similarly, in the deductive case, getting the hang of modus ponens or reductio ad absurdum, say, involves some understanding of the validity of these rules; extending one's knowledge on the basis of arguments of

---

<sup>2</sup> The terminology of 'causation in virtue of rationalization' is derived from Louise Antony (1989, pp. 167-8), who in turn acknowledges Joe Levine. It is as yet just a place-holder, of course, for what I regard as a crucial phenomenon central to a person's motivation in thought and action by reason quite generally.

these forms involves acquiring beliefs on the basis of this sense of how they must therefore be true, of why one is right to do so. It is notoriously difficult, though, to spell out precisely the manner in which acquiring and exercising cognitive capacities of this kind involve grasping and applying the norms governing them: what could grasping these norms possibly be, such that it is a plausible product of our education in the relevant area and appropriately explains future performance? (See Wittgenstein, 1958a, §§135-242; 1958b, pp. 89-100; and 1978, Part VI.)

## **2. Explicit Knowledge**

One way to envisage learning in these cases is as the explicit formulation either of necessary and sufficient conditions for the correct application of the term or of an abstract general statement of the inference-rule in question. This suggestion is wildly implausible. We do not apply the predicate ‘*x* is red’ by appeal to explicit necessary and sufficient conditions for its correct use; indeed, it is far from clear what non-circular such conditions would be. Similarly, we do not normally come to know the conclusion of a simple deductive argument by reflection on the general validity of arguments of that form.<sup>3</sup> More importantly, the idea that possession of these capacities consists in explicit knowledge of a rule for their correct exercise on every occasion is in any case open to the charge of vicious regress. For, very crudely, precisely the problematic relation between a person’s training and future rule-governed performance, which is supposed to be explained by her explicit formulation and consultation of the general rule, obtains either between her training and her formulation of the rule, or between her consultation of it and obeying it on some particular occasion.

---

<sup>3</sup> See Carroll (1895) for a revealing parody which points towards the difficulty in principle with the approach currently under consideration in the deductive case.

The fundamental objection here can be put in the form of a dilemma. Trainees experience a series of correct, and perhaps some incorrect, uses of the capacity in question. The explicit rule they supposedly thereby acquire is either to be construed as applicable directly, without interpretation, in every possible situation, or it is not. Suppose that it is. Then there are two problems. First, nothing in a person's training uniquely determines which particular such rule is to be associated with the practice in question. For example, our finite experience with arithmetic is equally compatible with rules for the operation denoted by '+' corresponding to standard addition, 'plus', and to any number of quaddition-like operations (Kripke, 1982), defined as follows:  $x$  quus  $y$  is equal to  $x$  plus  $y$  for all  $x$  and  $y$  except  $x=a$  and  $y=b$ ;  $a$  quus  $b$  is equal to  $c$ , which is not equal to  $a$  plus  $b$ , where  $a$  and  $b$  are large numbers not previously encountered in any addition/quaddition calculation. Second, a person's ability to bring the right such rule to mind in aid of directing her exercise of the cognitive capacity in question on some particular occasion is itself an ability of precisely the kind we are currently supposing requires explanation by appeal to a further explicit rule. So we have a regress.

The latter difficulty is nicely illustrated by an example of Wittgenstein's (1958b, p. 3). He considers a person's understanding of and compliance with the order to fetch a red flower from the meadow. The present account regards her understanding of the word 'red' in this order as consisting in her ability to bring to mind an explicit rule from which her selection of a correctly coloured flower can be read off directly, without any interpretation: an image of a red patch, say (although even this might well be thought, at least in principle, to require interpretation). But then how does the account handle a similar request for her to bring the right rule to mind, to imagine a red patch, say? Insofar as there is a puzzle how her previous learning experience enables her to comply with the initial order, there is equally a puzzle in this second case. Hence the current account of the first case, which simply

assumes her ability to summon up the correct colour patch at will in her imagination, fails to give a satisfactory elucidation of the capacity in question. Precisely the phenomenon to be explained features in the deliberate bringing to mind of the appropriate explicit rule or sign, consultation of which is supposed to constitute its explanation.

Alternatively - and this constitutes the second horn of the dilemma facing attempts to construe the grasp of norms involved in the acquisition and exercise of cognitive capacities in terms of (something like) explicit knowledge of necessary and sufficient conditions for correct application - one might construe such general rules other than as applicable directly in every case, that is to say, as standing in need of an interpretation to determine what is to be done when they come to be applied. Such a rule is therefore in itself open to any number of distinct interpretations prescribing incompatible actions in certain circumstances. Alone it stands to correct use just as the learning exposure to particular instances does. So the connection between the explicit rule and a person's exercise of the cognitive capacity in question requires precisely the kind of explanation that appeal to the rule on its own is supposed to provide. But then insisting that the bridge between a rule of this kind and its application in practice is to be made by an interpretation takes us straight back to the dilemma above. Either an interpretation can be applied directly, without further interpretation, in every case, or it cannot. In the first case, the relation we are simply assuming between grasp of the explicit rule and grasp of its correct interpretation is of precisely the kind to be explained. In the second case, the interpretation stands in need of further interpretation if it is to be of any use in determining actual behaviour on any particular occasion. Either way, we are launched on a vicious regress.

### **3. Blind Inclination**

An influential alternative to this whole idea of explicit formulation and consultation of abstract general norms of correctness in the acquisition and exercise of cognitive capacities, is the suggestion that a person is simply so configured, dispositionally speaking, by her training that she does, on the whole, go on correctly in the future: that is all her acquisition and resultant exercise of the capacity in question consists in. This is, in effect, to ignore completely the intuition that acquiring and exercising cognitive capacities each involve at least some appreciation of the norms governing them, relative to which one's performance is assessable as correct or incorrect. Yet to the extent to which we are right to regard this future performance as the purposive activity of a rational agent, justifiable in the light of these norms of correctness governing the practice in question, then I think that the initial intuition is sound, and any purely mechanical dispositional account of the matter is unacceptable.

To start to see why, consider what is involved in following a valid deductive argument with real understanding, in a way that yields knowledge of its conclusion. Of course there is always the possibility of a slip or unnoticed error, but in cases where I do not go wrong, I am correctly compelled by the argument to believe its conclusion. In following and fully understanding the argument, this compulsion is not simply a blind and mysterious manipulation of my beliefs by some reliable mechanism, however well established by evolution, benevolent hypnosis or whatever. I am not just a machine which runs along those rails. For if my following the argument is really to extend my knowledge, then my understanding of it must give me some appreciation of why I am right in believing its conclusion. I have to have some grip on how I thereby know the conclusion, and my belief should be guided by this understanding. A disposition to take beliefs on board in parallel with the steps of the argument is on its own insufficient for the argument to provide me with genuine knowledge. For such beliefs would come as a succession of mere hunches, wholly unsubstantiated *for me* by the de facto validity of the argument propelling my



endorsement of them.

Similarly, suppose a person knows that there are five boys in a room, and a hypnotist brings it about that regardless of the content of her next belief, it will produce a further belief that there are twelve children in the room. She sees that there are seven girls in the room, and forms the above belief as a result (see Lennon, 1990, p. 38). In parallel with Davidson's climber, her arriving at the final belief is causally dependent upon her having what are in fact reasons for that very belief, but not in the right way to constitute this a new piece of knowledge. The appropriate reason-giving relation is causally irrelevant, in the sense that her final belief is formed in complete ignorance of why she is right to hold it. Even the permanent attendance of such a benevolent hypnotist, reliably installing true beliefs, would do nothing on its own to rectify *this* ignorance, and so provides an inadequate model for possession of the cognitive capacity for addition. Furthermore, any eventual habituation she might attain to this peculiar situation would produce at best an ungrounded absence of puzzlement rather than any real appreciation of how she is onto the truth. However reliable or familiar, brute arithmetical inclinations along these lines remain *blind* hunches.

There is more to grasping the laws of logic or mathematical argument than simply being disposed to have one's beliefs mirror the moves they prescribe. Epistemologically productive reasoning is not a merely mechanical manipulation of belief, but a compulsion in thought *by reason*, and as such involves some conscious understanding of why one is right in one's conclusions. What drives one in genuine cases is precisely what one lacks if one has simply been drilled by the dictators of the International Academy of Logic, however benevolent they may be, to reason in ways for which one sees no evident rationale, no point or purpose (Dummett, 1991, ch. 9).

More generally, the problem for any purely mechanical dispositional account is that it is bound to ignore this sense of why one is right in exercising the capacity in question as one does, which is crucial if this is to make *cognitive* contact with the truth in the relevant area. It is the fact that such a capacity has as its point ascertaining the truth on some matter, the colour or shape of a perceived object, for example, or the conclusion of an a priori argument, which sets the norms for its correct exercise. Some appreciation of how what one is up to is onto this truth, is sensitive to the resultant norms, is therefore essential if exercising it is to be more than a blind mirroring of the norms, extrinsically, and in this sense only incidentally, in contact with the truth. If one's activity is really to serve as a source of knowledge and understanding, then one must have some grasp of the correctness of what one does. No brute disposition on the basis of one's training, simply and mechanically to go on in one direction rather than another in future, is suitable or sufficient on its own to constitute a genuinely cognitive capacity.

This difficulty is clearly related to those Michael Dummett stresses for any purely practical account of the knowledge constitutive of linguistic understanding: "to regard the understanding of a word or an expression purely as a practical ability is to render mysterious our capacity to know [directly] whether we understand" (1991, p. 93). We generally answer the question whether we have a certain practical ability either by induction from past success or by a present attempt to exercise it. But a person who lacks understanding cannot even try to exercise it, because he does not know what it *is* to try; and a person with it normally knows he understands without the need for either induction or a test. The account under consideration attempts to assimilate the cognitive to the purely practical. And the evident difference between them in directness of self-ascription is surely bound up with the essential involvement of a person's grasp of why what he is doing is correct in his possession of a cognitive capacity like linguistic understanding.

If I am right that the objection to a purely dispositional account of the acquisition and exercise of cognitive capacities is this failure to account for the subject's awareness of meeting the norms governing his practice, then it should be fairly clear that simply setting his blind inclinations in the context of other people with like practical abilities, generally agreeing in behaviour of the relevant kind, is powerless to improve matters. This is fatal to the surprisingly popular 'community view' of rule-following inspired by Kripke's (1982) 'sceptical solution' to what he regards as Wittgenstein's 'sceptical paradox'. It is very likely in practice that education by and communication with others are of great importance in cultivating an individual's sense of what he is up to and why he is right in exercising various cognitive capacities as he does. But simply adding to a single creature, which is mechanically disposed to behave in certain ways, a number of others with like inclinations cannot possibly constitute such comprehension in itself.

It should be equally clear that a composite view combining explicit knowledge with blind inclination is also unsatisfactory. The idea would be that rule-following is the joint product of (1) a brute causal connection between past and future behaviour, and (2) a theoretical grasp of an explicit rule for the practice in question. But this faces all the difficulties we have already seen for its individual components. First, it is simply a datum that most people have little, if any, reflective access to a general algorithm for the correct exercise of cognitive capacities they nevertheless uncontroversially possess. Second, the claim that any such rule uniquely determines the correct course of action in every case, in absolute isolation from the subject's inclination to go on in some particular way or another, is what generates the vicious regress we saw above in section 2. Third, and perhaps most important for my purposes, from the point of view of the subject herself, the inclination to apply a term in one way rather than another on a particular occasion, or to proceed to the

conclusion of a deductive argument, is, on this composite view, quite mysterious, an epistemologically ungrounded blind hunch. What is required for linguistic *understanding* or deductive *knowledge* is an awareness of why what one is doing in exercising it on some particular occasion is correct, how one is onto the relevant truth, which itself enters into the causal explanation of one's doing as one does: causation in virtue of rationalization, or compulsion *by reason*.

Bringing together the main elements of the discussion in this and the previous section, then, we can say that it is essential to a person's possession of a cognitive capacity of the kind we are concerned with that she have some coherent grasp of the point and purpose of her activity in exercising it, over and above her mere inclination to use it appropriately, perhaps in accord with others around her, which motivates her correct use of it on particular occasions and makes sense for her of why she is right in doing as she does. Nevertheless, this cannot consist in any explicit knowledge of an abstract general algorithm for its correct use, or full-blown reflective understanding of the norms governing the practice in question, independent of her actual behaviour in conforming to it. The problem, of course, is how to characterize this crucial mode of awareness.<sup>4</sup>

---

<sup>4</sup> I should emphasize the epistemological nature of my concern with rule-following. Wittgenstein (1958a, §§135-242) undermines two conceptions of the norms governing cognitive capacities: (1) the Platonic conception of infinitely extended rails defining correctness on every possible occasion quite independently of a competent rule-follower's inclination to go on, to which he is attached by explicit knowledge of their course at every point; and (2) the idea that they might be reduced to any actual user's isolated dispositions. Wright (1981) claims that these norms are constituted by *communal* inclinations in exercising the capacity in question. Justification for individual subjects' following the rule as they do is then to be given anti-realistically, in terms of assertibility rather than conformity with investigation-independent facts. McDowell (1984) takes the existence of realistic truth-conditions for granted, and argues that rule-following consists in finding it 'second nature' to go on in accord with them, without the need for justification from any independent codification of Platonic norms: indeed, he insists, the only available conception of the truth in the relevant area has to be given through the eyes of a competent person, in terms of how he finds it natural to proceed. In contrast with both of these, I aim to provide an account of justification which nevertheless preserves realism. I assume a spatial understanding of objective physical reality, and propose an account of what a person's grasp of the correctness of his judgements consists in in terms of his

#### 4. Compulsion by Reason

We started with the idea that mental causation has both causal and rationalizing aspects; and we saw in the case of intentional action that these are essentially united: the causal dependence of a person's acting upon her having the reason for which she acts obtains *in virtue of* the reason-giving relation between them. Sections 2 and 3 together build on this point in two equally important cases, of language use and deductive reasoning, and demonstrate the untenability of any account on which the two aspects of mental causation are treated independently. In particular, it is a mistake to construe the rationalizing aspect in terms of a person's reflective formulation and consultation of explicit general rules determining correct behaviour in every case; and it is equally unacceptable to regard the causal aspect as a purely mechanical blind inclination, related, if at all, only extrinsically and beyond the ken of the subject to the norms of correctness governing the practice in question.

The assumption behind the idea that we must in the end adopt whichever is supposed to be the lesser of these two evils is that physical mechanism provides the only model for causal dependence,<sup>5</sup> in the sense that any *justification* is bound to be independent of the particular process in question, derived rather from an abstract general rule, from which the correctness of the occurrence at issue follows by subsumption. Given this, the familiar debate then rages over whether this extrinsic 'reason' has to be *cognitively accessible* to the subject (however precisely that notion is to be spelt out). Insisting that it does generates the regress made vivid by Wittgenstein's rule-following considerations (section 2. above). Denying that it does

---

egocentric perceptual perspective on this world, which resists both untenable appeal to a general algorithm, and anti-realistic reduction of truth to communal consensus. At least, a sketch of how this might be done is what I offer in the following section.  
<sup>5</sup> For elucidation of and objection to this restrictive view, see Strawson (1985) and Child (1994, chs. 3 & 6).

leads to the blind reliabilism rejected in section 3. Yet these seem to be the only options.

Obviously the assumption responsible for this dilemma is at fault. What is required is an account of the causal dependence in question which itself makes essential appeal to some grasp on the part of the subject of the point of what she is doing and why she is right in doing it. This intrinsic and motivating grasp of the reasonableness of her behaviour, which makes the causal dependence one in virtue of rationalization, should nevertheless be quite different from any independent reflective understanding or explicit knowledge of a general rule governing the practice in question, from which the appropriate action is to be inferred on any particular occasion. Exactly how causal dependence in virtue of rationalization along these lines should be elucidated in each case in which it is operative is a difficult and in my view extremely important matter. I shall sum up the position at which we have arrived and then end this section with an illustration of the general approach I am proposing in connection with a case we have not yet discussed: the dependence of a person's present beliefs about the way things are in the external world around him upon his conscious perceptual experience. The point of this is to show that there is at least room for a view of this form.<sup>6</sup>

Correctly applying a term or following a deductive argument, with real understanding, involves a kind of awareness of what one is up to and why one is right in doing as one does. Learning to do these things involves coming to see one's activity in this way. This is far more than any epistemologically insignificant blind inclination to go on in one way or another; yet it need not, and in the most basic cases cannot, consist in the acquisition and interpretation of an explicit general rule. What

---

<sup>6</sup> Lucas (1955) has ingenious further demonstration of the possibility of what he calls 'singular reasons', combining genuine normativity with the absence of any universal principle under which they are subsumed.

is required is rather a model of rational motivation and justification which is neither mechanically naturalized away to nothing nor dependent upon any reflective detachment from the actual activities of acquiring and exercising cognitive capacities to a point from which a completely disengaged, universal theoretical justification can be given. This mode of awareness can, I think, be illustrated by the role of a person's direct conscious perception of the way things are 'out there' in explaining his current beliefs about the external world around him. Indeed, I think it is a peculiarly perceptual mode of awareness, even in the cases of a person's grasp of how he is getting things right in applying a term as he does or in being influenced in his beliefs by argument. It lends itself very naturally to perceptual metaphor because it is connected with consciousness in a way that resists capture in terms of abstract rational relations between purely propositional contents: it is a matter of *seeing* why one is right in doing as one does.<sup>7</sup> So I shall now expand a little on the paradigm perceptual case.

Perception, like reasoning, and linguistic judgement, is a mode of belief acquisition. It constitutes a cognitive capacity of the kind we have been discussing, insofar as it serves its subject as a source of knowledge about how things are in the external world around him. Philosophical accounts of its undeniable status as such standardly fall into one of two broad categories, which I shall call *reflective internalism* and *reliabilism*. Theories of the first kind insist that the epistemic status of perceptual belief depends upon the subject's possession (or possible possession) of

---

<sup>7</sup> Thus, causation in virtue of rationalization essentially introduces the idea of a person's perspective on the world, or subjective point of view. It is therefore quite different from what might be called 'causation in virtue of gravity', say, where something that happens to a given massive body is explained as a particular instance of a perfectly objective pattern in the way such things generally tend to behave, grasp of which is possible without the introduction of any notion of a personal perspective. For useful further discussion of this contrast, see McDowell (1985) and Campbell (1993). The point I am making here is crucial to understanding the simulation theorist's case against the theory theory of mind. See the symposium on 'Understanding the Mental: Theory or Simulation?' in Peacocke (1994) for an introduction to this debate.

a justified belief in the general reliability of perception, which constitutes his essential understanding of why he is right in believing the world to be as it is presented as being in perception. Epistemic warrant, on this account, has an inferential structure, in which the conclusion that things are thus and so is derived (or is at least derivable in principle) by the perceiver from two independent premises, each with its own quite separate justification: (a) things seem to be thus and so (perceptually); and (b) the way things seem to be (perceptually) is a reliable guide to the way they are. Reliabilist theories, on the other hand, place all the responsibility for the epistemic status of perceptual belief upon the de facto reliability of perception: if a true belief is arrived at by a method (however these are to be individuated) which normally (however this is to be spelt out) yields the truth on such matters, then it is a case of knowledge. Perception constitutes a source of knowledge precisely insofar as it is reliable in this sense.

Adopting terms devised by Wilfrid Sellars (1963), the fundamental difference between the two broad approaches can be put as follows. Let us say that a belief has *authority* if it is the product of the operation of a mode of belief formation which is generally reliable on such matters. On the reliabilist account, a true belief constitutes knowledge precisely if it has such authority. The internalist intuition, on the other hand, is that epistemic justification requires a further *recognition* of this authority, in some sense, *by the subject*. Very much as we have just seen in the case of rule-following, though, a crucial assumption is involved in the idea that the only alternative to crude reliabilism is the reflective internalism I define above. This is the assumption that the only model for a person's recognition of the authority of a certain belief is his explicit knowledge of the general reliability of beliefs arrived at in that way on such matters, where this is knowledge which is supposed to be quite independent of his actual acquisition of the relevant belief on the particular occasion in question. Again this is an instance of the misguided attempt to treat the causal and



reason-giving aspects of epistemic warrant as independent and independently intelligible. Given this assumption, the internalist is stuck with an abstract, inferential conception of justification by the subject's appeal to some extrinsically warranted belief in the general reliability of perception, which faces all the familiar difficulties: either those of the foundationalist, both in characterizing the content and epistemic standing of a person's introspective knowledge of the way things perceptually seem to him to be, and in giving a satisfactory account of his knowledge of the general reliability of perceptual belief; or those of the coherentist, both in avoiding unacceptable circularity, and in anchoring perceptual knowledge to the actual truth rather than that in some other merely possible world.<sup>8</sup> On the other hand, the reliabilist's purely mechanical conception of the causal dependence of perceptual belief upon the way things are in the world around the perceiver fails to avoid an untenable assimilation of genuine knowledge to blind hunch or benevolent brain-washing, by ignoring the essential involvement in the former of some grasp on the part of the subject of why he is right in believing what he does, and how he thereby knows the way things are around him.

What is required, as we have already seen in general terms above, is a rejection of the detachment of an independent reflective justification for perceptual belief from some underlying mechanical process of belief formation, in a context which nevertheless respects the need for a person to recognize the authority of perception *in some sense* if this is to constitute a genuine source of knowledge. We need an account of the dependence of perceptual belief upon experience in which causal and rationalizing or justificatory elements are integrated aspects of a single phenomenon. I believe that this can be given in terms of the perspectival, egocentric spatial content of perceptual experience. In veridical perception, a person is presented

---

<sup>8</sup> See Haak (1993, chs. 1-3) for a helpful survey and development of these standard objections.

with the way things are around him. They are displayed as being thus and so *around him*, in the sense that they are given as being in certain spatial relations with him. He sees the tree *in front*, feels something hard *over to the right*, or hears the bang *behind*. Although he does not figure explicitly as one object among many in his perception, it is nevertheless the spatial relations between the things he does perceive in this way *and him* which determine whether or not it is veridical. Furthermore, and partially explanatory of the peculiarity just noted, these perceived egocentric relations have their spatial significance for him, not in terms of his abstract theoretical understanding of what it is for any object to stand in the relevant spatial relations to any other, into which he feeds himself as one of the related objects, but rather directly in terms of his appreciation of the consequences of their obtaining for his own action upon them and *perception of them*. In particular, then, the significance for the subject of such perceptually presented egocentric spatial relations is given in part in terms of the enabling conditions on his perception of the very things standing in them. For example, to see something moving behind an obstacle is to see it moving out of sight, in such a way that such and such relative movements (of perceiver, obstacle and object) would, other things being equal, bring it back into view. So things are presented in perception as being *accessibly* thus and so to the perceiver in that very experience, in virtue of the way in which their spatial relations with him are presented in experience.

Four key features of this account should be stressed in the current context. First, perception presents its objects as determinately spatially related to the perceiver. Second, that certain objects are thus and so spatially related to a person can be displayed in different ways; and we can distinguish between such ways in terms of what it means for the person to have things displayed as so-related in each of them. Third, the significance to a perceiver of things' being displayed as standing in certain spatial relations with him in the peculiarly perspectival way this occurs in his direct

perceptual experience of them is given in part in terms of the systematic dependence of their perceptibility upon such spatial relations. So, fourth, a perceptual presentation of something as some way at some egocentric location is a presentation of its being perceptibly that way in that very experience.

In perceiving things to be thus and so, therefore, a person grasps how he is thereby onto the way things actually are in the world around him: in virtue of the spatial relations things stand in to him, which enable his perception of the way they are. Thus, in judging things to be that way, his perceptual belief is causally dependent upon his appreciation in experience of why he is right to hold it.<sup>9</sup> A person's direct perception of the way things are in the world around him provides a paradigm case of the integration of causation and rationalization constitutive of compulsion in thought by reason, which is essential, quite generally, if the capacities we have been concerned with, for language use, deductive reasoning, and now perception, are to make a genuinely cognitive impact on their subject, in the growth of his knowledge and understanding. The next move, and one I cannot possibly even start on here, would be to attempt to exploit the key features of the perceptual paradigm in giving an illuminating characterization of the causal dependence in virtue of rationalization which obtains between the exercise and the possession of these other cognitive capacities that we have been discussing.

## **5. Conclusion**

Linguistic understanding, deductive reasoning and perception are all cognitive capacities, in that they serve in some way for their subject as a source of knowledge. On any occasion of their successful use, a person's acquisition of knowledge causally

---

<sup>9</sup> My (forthcoming) constitutes an elaboration and defence of this account of perceptual knowledge.

depends upon his possession and exercise of the capacity in question. Arthur's knowledge that Betty is clever depends upon his understanding the word 'clever'; his deductively-based knowledge that  $p$  depends upon his grasp of the validity of inference rules instantiated in the argument; and his perceptual knowledge of how things currently stand directly in front of him depends upon his current perceptual experience of them. These dependencies are instances of mental causation. Not only are they causal, but they also have a rationalizing aspect: given his acquaintance with her, Arthur's understanding of 'x is clever' gives him a reason to apply it to Betty; given his belief in the relevant premises, his grasp of modus ponens, or whatever, gives him a reason to believe that  $p$ ; and in the absence of reason to doubt, his perceptual experience gives him a reason to believe that things are thus and so in front of him. Neither the causal nor the rationalizing aspect of these dependencies is correctly captured by regarding these as two independent and independently intelligible components. The idea that the rationalization necessarily consists in Arthur's appeal to an independent universal principle in justification of his exercise of the capacity in question on any particular occasion by subsumption launches a vicious regress; whereas a purely mechanical conception of the causation involved inevitably fails to provide him with any grasp of what he is up to in exercising the relevant capacity, and why he is right in doing so, which is essential if this is to make any genuinely cognitive contact with the facts putatively thereby *known*. The two are aspects of a single *sui generis* phenomenon, causal dependence in virtue of rationalization, which is paradigmatically illustrated in the perceptual case, by the dependence of the way a perceiver takes things to be around him upon their being displayed that way to him in experience, in a way which at the same time makes immediately evident the accessibility of their being that way to him in that very experience, how he thereby knows the way things are around him, and why he is right in taking them to be the way he does.<sup>10</sup>

---

<sup>10</sup> Many thanks to John Campbell, David Charles, Bill Child, Naomi Eilan, Kathleen

## References

- Antony, L. 1989. 'Anomalous Monism and the Problem of Explanatory Force',  
*Philosophical Review*, **98**, 153-187.
- Brewer, B. Forthcoming. 'The Foundations of Perceptual Knowledge'.
- Campbell, J. 1983. 'A Simple View of Colour'. In J. Haldane and C. Wright (eds.),  
*Reality, Representation and Projection*. Oxford: Clarendon Press
- Carroll, L. 1895. 'What the Tortoise said to Achilles'. *Mind*, **4**, 278-280.
- Child, W. 1994. *Causality, Interpretation and the Mind*. Oxford: Clarendon Press.
- Davidson, D. 1963. 'Actions, Reasons and Causes', *Journal of Philosophy*, **60**, 685-700.
- 1973. 'Freedom to Act'. In T. Honderich (ed.), *Essays on Freedom of Action*.  
London: Routledge.
- Dummett, M. 1991. *The Logical Basis of Metaphysics*. London: Duckworth.
- Haak, S. 1993. *Evidence and Inquiry*. Oxford: Basil Blackwell.
- Kripke, S. 1982. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Lennon, K. 1990. *Explaining Human Action*. London: Duckworth.
- Lucas, J. R. 1955. 'The Lesbian Rule', *Journal of Philosophy*, **30**, 195-213.
- Peacocke, C. 1994. *Objectivity, Simulation and the Unity of Consciousness*:  
*Proceeding of the British Academy*, 83. Oxford: Clarendon Press.
- McDowell, J. 1984. 'Wittgenstein on Following a Rule', *Synthese*, **58**, 325-363.
- 1985. 'Functionalism and Anomalous Monism'. In E. LePore and B.  
McLaughlin (eds.), *Actions and Events: Perspectives on the Philosophy of*

---

Lennon, Adrian Moore, Hugh Rice, Johannes Roessler, Helen Steward, Nick Taylor, Ralph Walker and Timothy Williamson for useful and enjoyable discussions of these issues and helpful comments on an earlier version of the paper. Thanks also to Susan Seville for speedy and accurate typing of a hand-written first draft.

- Donald Davidson. Oxford: Blackwell.
- Sellars, W. 1963. 'Empiricism and the Philosophy of Mind'. In his *Science, Perception and Reality*. London: Routledge.
- Stoutland, F. 1976. 'The Causation of Behaviour', *Acta Philosophica Fennica: Essays on Wittgenstein in Honour of G. H. von Wright*, **28**, 286-325.
- Strawson, P. F. 1985. 'Causation and Explanation'. In B. Vermazen and M. Hintikka (eds.), *Essays on Davidson: Actions and Events*. Oxford: Clarendon Press.
- Taylor, C. 1964. *The Explanation of Behaviour*. London: Routledge.
- Wittgenstein, L. 1958a. *Philosophical Investigations*. Oxford: Blackwell.
- 1958b. *The Blue and Brown Books*. Oxford: Blackwell.
- 1978. *Remarks on the Foundations of Mathematics*, 3rd edn. Oxford: Blackwell.
- Wright, C. 1981. 'Rule-Following, Objectivity and the Theory of Meaning'. In S. H. Holtzman and C. M. Leich (eds.), *Wittgenstein: To Follow a Rule*. London: Routledge.

## ***Mental Causation: Compulsion by Reason***

**BILL BREWER**

I'll try to do two things: first, sketch the main line of argument from my paper, although this cannot make it wholly perspicuous to those who haven't read it; and second, make some comments about Tim's paper, and the relations between his views and mine.

### **1. My main argument**

Mental causation I take to be, centrally, that which figures in reason-giving explanation. When we give a person's reason for believing or doing something, we cite a cause of her forming or retaining the belief, or performing the action in question, which makes this behaviour intelligible to us as rational given her other attitudes and circumstances. Furthermore, if this is to be genuinely explanatory, then the reason-giving relation itself must be causally relevant. I am interested in the nature and commitments of this reason-giving model of causal explanation. Since it seems to me that we have more of a grip on the notion of believing something for a reason than we do on that of acting for a reason, through its direct connection with epistemological questions of knowledge and justification, I focus on this: for the main part of the paper in connection with language use and deduction.

Linguistic understanding and deductive reasoning enable our acquisition of knowledge. Their successful application is a matter of making appropriate judgements in the light of some conception of why these are right: judgements which causally depend upon our linguistic understanding or grasp of the relevant inference

rules in virtue of the reasons these give us for them. But what exactly does this involve?

I argue against two influential models of this causal dependence in virtue of rationalization, minor variations on which are often presumed to exhaust the field.

First, it cannot consist in making a judgement on the basis of explicit knowledge of a general rule, for applying the term in question, or expressing the valid pattern of inference involved. For this is bound to lead to a regress. Either the rule can be applied without interpretation in every case or it cannot. If it can, then another is required to explain one's knowledge of the right such rule. If it cannot, then another is required to explain one's correct application of it in any particular case.

Second, the reason-giving causal dependence of our knowledge upon our linguistic understanding or grasp of certain basic rules of inference cannot simply consist in its being the product of a general disposition to use the relevant expressions in appropriate circumstances or to arrive at conclusions licensed by the inference rules in question, without any grasp whatsoever of what we are upto in doing so, or why we are right in doing it. However reliable they may be, such deliverances are just blind hunches, for which we have no rational epistemic justification.

Furthermore, setting such dispositions in the context of a community of other people who share them does nothing to rectify this; and a mixed model combining reliable but blind inclinations with explicit knowledge of a general justification for them is also untenable, as this ultimately inherits the difficulties of each of its components.



The impression that these effectively exhaust the options rests on an assumption I reject: that physical mechanism provides the single paradigm for causal dependence. On this view, the justification for any judgement a person makes will be independent of the particular process by which it is formed, and given rather by a general rule from which the correctness of this process follows by subsumption. Given that, the only remaining question is whether this general rule has to be known, or at least knowable, by the subject, or whether it simply has to be true. If it has to be cognitively accessible in some way, then we have a version of the first model above and face the Wittgensteinian regress. If it has simply to be true then we have an untenable form of blind reliabilism.

What is required is an alternative paradigm for causal dependence, which already makes essential appeal to a person's grasp of the point of what she is doing and why she is right in doing it: a paradigm provided directly by the rationalizing explanations of commonsense psychology, rather than to be constructed by adding hopefully to the blind mechanism of physical explanation.

I finish the paper with a sketch of how my central line of argument unfolds in another very important case of mental causation: the dependence of a person's beliefs about the world around him on his current perceptual experience. In doing so, I hope at least to make it plausible that the required phenomenon of intrinsically rationalizing casual dependence exists and can fruitfully be investigated and elucidated in a way which avoids the pitfalls of the two standard models.

Perception is a source of knowledge about the external world. It is often regarded simply as a mechanism for the acquisition of beliefs about the way things are out there whose epistemic status depends in some way upon the general rule that it is reliable on such matters. The big issue is then supposed to be exactly how this

dependence should be characterized. At the two extremes in this debate are those who regard the truth of the rule - de facto reliability - as sufficient for the status as knowledge of any true belief acquired through perception; and those who insist that the rule of reliability must itself be known. On the former, pure reliabilist view, the perceiver has no conception how he might be right in believing what he does, which offends against the internalist intuitions of many including myself. Yet the latter view is famously subject to scepticism or circularity.

The choice only arises on the assumption that the process of belief acquisition itself involves no sense on the subject's part of how he is onto the truth in believing things are as he takes them to be around him on the basis of perception. This understanding what he is upto is therefore supposed either to be unnecessary for perceptual knowledge, or to be provided by his reflective knowledge of the general reliability of the ground floor mechanism. But that debate would be irrelevant if the basic process of perceptual belief formation in itself were essentially to involve some grasp on his part of how he is thereby right about the world around him. I believe that this is the case, since the perspectival nature of perceptual experience presents the world as epistemically accessible to the subject. It does so, to cut a long story short, in virtue of its egocentric spatial content: the significance to the perceiver of things being where they are presented as being, when they are perceptually presented as being there is given directly in terms of the consequences of their being there for his perception of them. They are therefore presented as perceptibly thus and so in that very experience.

The key to understanding mental causation is to recognize its causal and reason-giving aspects as integrated in an essentially rationalizing mode of causal dependence, rather than independent elements of a composite of mechanical

causation and reflective justification by subsumption under a general rule - either known by the subject or simply true.

To add even more to the speculative nature of what I have to say here, let me end the first part of my presentation by proposing the following equivalence.

$S$  is conscious that  $p \equiv S$  has a non-inferential reason, or right, to judge that  $p$

The task in any area in which such basic epistemic rights exist, is simultaneously to solve for what it is for a person to have such a right and for the nature of the appropriate mode of consciousness.

## **2. Comments on Crane**

It may appear that there is little relation between what we Tim and I have to say. And I apologise for making direct engagement more difficult by not responding explicitly to his paper in mine; but I didn't get it in time for there to be any chance of doing this. Nevertheless, I do think there are important points of contact, and tension. I also have some independent comments about his central line of argument.

As I see it, the basic point of Tim's paper is to bring out a contradiction between endorsing the premises of a general argument for identity versions of physicalism, with the dominant physicalist rejection of mental-physical identities. Insofar as the argument is valid, this is obviously inconsistent. But he wants to extract far more than this: the conclusion that non-identity-theory versions of physicalism are either unmotivated or unstable. To do this, he assumes (a) that this form of argument is the only possible motivation for anything that might be called 'physicalism'; and (b) that only a non-physicalist can resist its premises. I am far from convinced by

either of these. Of course it matters to get straight about what physicalism involves; and in Tim's formulation - "the mental is physical" - this amounts to the question what readings of 'is' are permitted. The more restrictive he is here, the more plausible his assumptions, but the less significant his conclusion. The more permissive, the less plausible the assumptions, but the more impressive his conclusion if it were acceptable. E.g. if it is necessarily the 'is' of strict identity, then the assumptions are pretty reasonable, but the conclusion is empty; if it allows a much looser reading of some asymmetric dependence of the mental upon the physical, then although the conclusion would be substantive, the required assumptions are extremely tendentious.

I cannot survey all the possibilities here and make good this general worry. But I shall try to sketch a view which may be worth calling physicalist and which seems to me not obviously to succumb to Tim's accusation of being either unmotivated or unstable.

First, it is motivated other than by an argument of Tim's basic form - unless the mental is physical, mental causation is incompatible with the completeness of physics. Cartesian dualism is incoherent. The only persisting substances there are are space-filling material things whose mental properties, if any, supervene upon their physical properties. But this supervenience is not brute, a kind of property parallelism set up by divine will. Hence the way a person is mentally in some way depends upon the way s/he is physically. This is what makes it intelligible how there can be mentality in a world simply of space-filling material things. In this sense, therefore, the mental is physical, although the 'is' here is nothing like that of strict identity.

Not only is this view - whether you want to call it physicalist or not - not motivated by arguments of the kind Tim outlines, but it is also perfectly consistent with rejecting them. Take Lewis' argument. Its first, a priori premise is that mental

properties are defined by their causal explanatory roles in connection with perceptions, actions and other mental states. Even given the second, empirical premise that physical phenomena have complete physical causal explanations, the required identities follow only on the assumption that the relevant perceptions and actions are themselves identical to certain physical phenomena. This is obviously question-begging. Similarly, the second, overdetermination argument only provides a motivation for identifying mental causes with internal physical events on the question-begging assumption that the relevant effects of the former - e.g. S's raising her arm or drinking a glass of water - are identical to certain physical effects of the latter.

Furthermore, and relatedly as Tim stresses, this view therefore avoids his problem of mental causation, since it simply rejects the idea that the two types of causation, mental and physical, are in competition with each other in a way which can only be resolved by identifying their relata - an idea which he rightly ascribes to an unacceptably physical conception of mental causation. On my view, different kinds of explanation give us different models of causal dependence. The mechanical action of one physical thing upon another - crudely, causation as impact - is only one such model. Folk psychology operates with another quite different model, of intrinsically rationalizing causal dependence; and its claims to causal relevance must be assessed in that context. The causal relevance of the mental is not dependent upon its being identified with some form of mechanical causal impact of the condition of a person's innards upon the physical movement of her body. Rather, it is a crucial component of a quite different type of explanation, which is essential to our making sense of what people, both ourselves and others, say and do.

This is very much in the spirit of Tim's endorsement of Burge's recent views. But he resists classifying his position as a rejection of the homogeneity of mental and

physical causation, which seems like the obvious way to think of it in his terms - I can't myself see what else that rejection would come to. Perhaps he doesn't want to call it that, since doing so, as he admits, threatens his conclusion that the problem of mental causation affects any form of physicalism which is not an identity theory. I think the line I've sketched precisely avoids this though, and I don't see anything in what he says which goes against it.

Tim begins with some bad arguments for the claim that mental causation is unintelligible unless mental phenomena are identical to physical phenomena. He admits that this is an untenable position. But then he hopes to recycle those same arguments to generate a problem explaining mental causation for any other conception of a dependence of the mental upon the physical. I don't see why, given these must be bad arguments in the first place, he thinks they are any better in their recycled form, unless he just helps himself to his conclusion by defining anyone who resists them as a non-physicalist.